

Realization of finite-state mixing Markov chain as a random walk subject to a synchronizing road coloring

Kouji Yano⁽¹⁾⁽²⁾ and Kenji Yasutomi⁽³⁾

August 31, 2010

Abstract

A mixing Markov chain is proved to be realized as a random walk in a directed graph subject to a synchronizing road coloring. The result ensures existence of appropriate random mappings in Propp–Wilson’s coupling from the past. The proof is based on the road coloring theorem. A necessary and sufficient condition for approximate preservation of entropies is also given.

Keywords and phrases: Markov chain, random walk in a directed graph, road coloring problem, Tsirelson’s equation, coupling from the past.

AMS 2010 subject classifications: Primary 60J10; secondary 37A35; 05C81; 37H10.

1 Introduction

Our purpose is to realize a mixing Markov chain as a suitable random walk in a directed graph, which is generated by a sequence of independent, identically distributed (*IID* for short) random variables taking values in the set of mappings of the state space.

1.1 Notations

Let V be a set of finite symbols, say $V = \{1, \dots, m\}$. Let $Y = (Y_k)_{k \in \mathbb{Z}}$ be a Markov chain taking values in V and indexed by \mathbb{Z} , the set of all integers. We call Y *stationary* if $(Y_{k+n})_{k \in \mathbb{Z}} \stackrel{d}{=} (Y_k)_{k \in \mathbb{Z}}$ for all $n \in \mathbb{Z}$. We write $Q = (q_{x,y})_{x,y \in V}$ for the one-step transition probability matrix of Y , i.e.,

$$q_{x,y} = P(Y_1 = y | Y_0 = x), \quad x, y \in V. \quad (1.1)$$

The n -th transition probability matrix is given by the n -th product $Q^n = (q_{x,y}^n)_{x,y \in V}$. We call Y *irreducible* if for any $x, y \in V$ there exists a positive number $n = n(x, y)$ such that $q_{x,y}^n > 0$. We call Y *aperiodic* if the greatest common divisor among $\{n \geq 1 : q_{x,x}^n > 0\}$ is one for all $x \in V$. We call Y *mixing* if Y is both irreducible and aperiodic, which is

⁽¹⁾ Graduate School of Science, Kobe University, Kobe, JAPAN.

⁽²⁾ The research of this author was supported by KAKENHI (20740060).

⁽³⁾ Department of Mathematical Sciences, Ritsumeikan University, Kusatsu, JAPAN.

equivalent to the condition that there exists a positive integer r such that $q_{x,y}^r > 0$ for all $x, y \in V$.

Let Σ denote the set of all mappings from V to itself. For $\sigma_1, \sigma_2 \in \Sigma$ and $x \in V$, we write $\sigma_2\sigma_1x$ simply for $\sigma_2(\sigma_1(x))$.

Definition 1.1. Let Q be the one-step transition probability matrix of a stationary Markov chain. A probability law μ on Σ is called a *mapping law* for Q if

$$q_{x,y} = \sum_{\sigma \in \Sigma: \sigma x = y} \mu(\sigma), \quad x, y \in V. \quad (1.2)$$

Definition 1.2. For a probability law μ on Σ , a μ -random walk is a Markov chain $(X, N) = (X_k, N_k)_{k \in \mathbb{Z}}$ taking values in $V \times \Sigma$ such that $N = (N_k)_{k \in \mathbb{Z}}$ is IID with common law μ such that each N_k is independent of $\sigma(X_j, N_j : j \leq k-1)$ and

$$X_k = N_k X_{k-1} \quad \text{a.s. for } k \in \mathbb{Z}. \quad (1.3)$$

Let $Y = (Y_k)_{k \in \mathbb{Z}}$ be a stationary Markov chain with one-step transition probability matrix Q . Let (X, N) be a μ -random walk. Then it is obvious that $Y \stackrel{d}{=} X$ if and only if μ is a mapping law for Q . For any stationary Markov chain Y , we can find a mapping law μ for Q (see Lemma 3.1).

Let us illustrate our notations. See Figure 1 below, where $V = \{1, 2, 3\}$ and

$$Q = \begin{bmatrix} q_{1,1} & q_{1,2} & q_{1,3} \\ q_{2,1} & q_{2,2} & q_{2,3} \\ q_{3,1} & q_{3,2} & q_{3,3} \end{bmatrix} = \begin{bmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{bmatrix}. \quad (1.4)$$

Let $\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}$ and $\sigma^{(4)}$ be elements of Σ characterized by Figures 2, 3, 4 and 5 below, respectively. The transition probability Q possesses several mapping laws; among others, we have $\mu^{(1)}$ and $\mu^{(2)}$ defined as follows:

$$\mu^{(1)}(\sigma^{(1)}) = \mu^{(1)}(\sigma^{(2)}) = \mu^{(1)}(\sigma^{(3)}) = 1/3, \quad (1.5)$$

$$\mu^{(2)}(\sigma^{(3)}) = 2/3, \quad \mu^{(2)}(\sigma^{(4)}) = 1/3. \quad (1.6)$$

The two random walks (X, N) corresponding to $\mu^{(1)}$ and $\mu^{(2)}$ have distinct joint laws, but have identical marginal law of X which is a Markov chain with one-step transition probability Q .

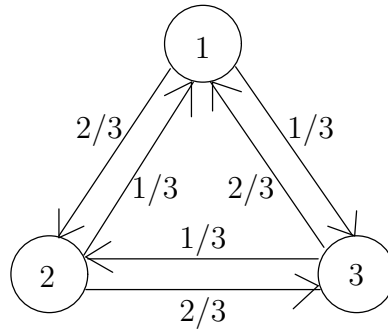


Figure 1. Transition probability

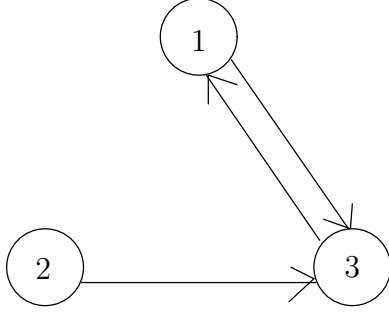


Figure 2. $\sigma^{(1)}$

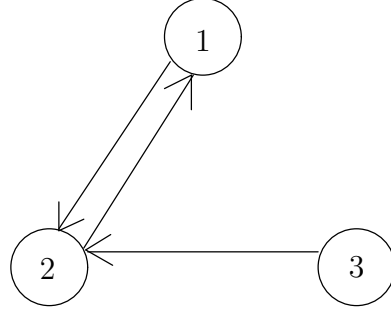


Figure 3. $\sigma^{(2)}$

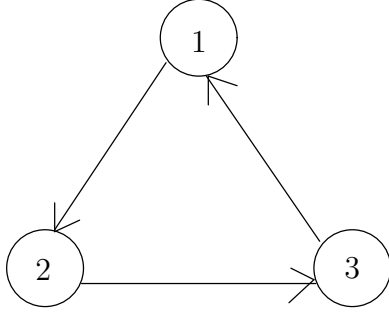


Figure 4. $\sigma^{(3)}$

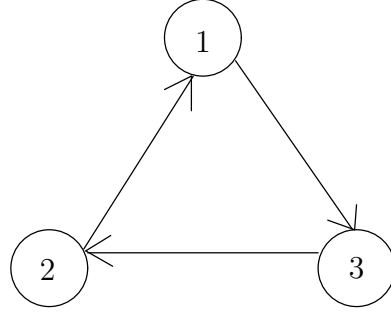


Figure 5. $\sigma^{(4)}$

1.2 Realization of mixing Markov chain as a μ -random walk

Our aim is to choose a mapping law μ which satisfies a nice property.

Definition 1.3. A subset Σ_0 of Σ is called *synchronizing* if there exists a sequence $s = (\sigma_p, \dots, \sigma_1)$ of elements of Σ_0 such that the composition product $\langle s \rangle := \sigma_p \cdots \sigma_1$ maps V onto a singleton.

First one of our main theorems is the following.

Theorem 1.4. Suppose that $Y = (Y_k)_{k \in \mathbb{Z}}$ is mixing. Then one can choose a mapping law μ for Q so that μ has synchronizing support.

Theorem 1.4 will be proved in Section 3.

Let us explain how our μ -random walk is related to road coloring. The support of μ , which we denote by $\{\sigma^{(1)}, \dots, \sigma^{(d)}\}$, induces the adjacency matrix A of a directed graph (V, A) which is of constant outdegree, i.e., from every site there are d roads laid. Then each element $\sigma^{(1)}, \dots, \sigma^{(d)}$ may be regarded as a road color so that no two roads from the same site have the same color. For a μ -random walk (X, N) , the process X moves in the directed graph (V, A) being driven by the randomly-chosen road colors indicated by N via equation (1.3). Thus we may call (X, N) a *random walk in a directed graph subject to a road coloring*. For example, the directed graphs induced by $\mu^{(1)}$ and $\mu^{(2)}$ which are defined in (1.5) and (1.6), respectively, are illustrated as Figures 6 and 7, respectively.

Since $\sigma^{(1)}\sigma^{(2)}V = \{3\}$, we see that the support of $\mu^{(1)}$ is synchronizing, while we can easily see that the support of $\mu^{(2)}$ is non-synchronizing.

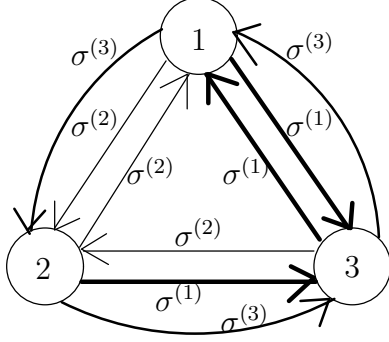


Figure 6. The graph induced by $\mu^{(1)}$

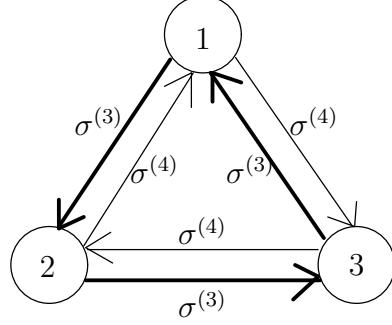


Figure 7. The graph induced by $\mu^{(2)}$

Let us come back to the general discussion. If (X, N) is a μ -random walk and if the support of μ is synchronizing, then the process X may be represented as

$$X_k = F(N_k, N_{k-1}, \dots), \quad k \in \mathbb{Z} \quad (1.7)$$

for some measurable function $F : \Sigma^{-\mathbb{N}} \rightarrow V$. In fact, define

$$T(N_j : j \leq k) = \max\{l \in \mathbb{Z}, l < k : N_k N_{k-1} \cdots N_l V \text{ is a singleton}\}, \quad (1.8)$$

where we follow the convention that $\max \emptyset = -\infty$. Since the support of μ is synchronizing, it holds that $T(N_k, N_{k-1}, \dots)$ is finite a.s. for all $k \in \mathbb{Z}$, so that we may define

$$X_k = N_k N_{k-1} \cdots N_{T(N_j : j \leq k)} x_0, \quad k \in \mathbb{Z} \quad (1.9)$$

for a fixed element $x_0 \in V$, but the resulting random walk does not depend on the choice of x_0 . This is such a representation as (1.7).

Letting $k = 0$ in the identity (1.9), we have

$$X_0 = N_0 N_{-1} \cdots N_{T(N_j : j \leq 0)} x_0. \quad (1.10)$$

This shows that the stationary law of the Markov chain may be simulated exactly from an IID sequence. This method was a central idea of *Propp–Wilson’s coupling from the past* ([13]; see also [9]). Our Theorem 1.4 assures theoretically that for any mixing Markov chain there always exists an appropriate mapping law such that Propp–Wilson’s algorithm terminates almost surely.

For the study of μ -random walks in the case of non-synchronizing supports, see Yano [17]. Equation (1.3) is called *Tsirelson’s equation in discrete time*; see Yor [20], Akahori–Uenishi–Yano [3], Yano–Takahashi [18], Yano–Yor [19] and Hirayama–Yano [10, 11] for the details.

The representation $Y \stackrel{d}{=} X = F(N)$ of Y by an IID sequence N of the form (1.7) is called a *non-anticipating representation*. Rosenblatt ([14] and [15]) obtained a necessary and sufficient condition for a Markov chain with countable state space to have a non-anticipating representation $Y \stackrel{d}{=} X = F(N)$ where $N = (N_k)_{k \in \mathbb{Z}}$ is an IID with uniform law on $[0, 1]$.

1.3 Condition for approximate preservation of entropies

Let Y and (X, N) as in Theorem 1.4. Let us compare the amounts of information of Y and N in terms of their entropies. See the standard textbook [4] for basic theory of entropies. Let λ be the stationary law of Y and define

$$h(Y) = - \sum_{x,y \in V} \lambda(x) q_{x,y} \log q_{x,y} \quad (1.11)$$

and

$$h(N) = - \sum_{\sigma \in \Sigma} \mu(\sigma) \log \mu(\sigma). \quad (1.12)$$

Since $Y \stackrel{d}{=} X$ and X is a measurable function of N as in (1.9), we have

$$h(Y) \leq h(N). \quad (1.13)$$

Note that Ornstein–Friedman’s theorem ([12] and [8]) asserts that two mixing Markov chains which have common entropy are isomorphic. By this theorem, we see that, if equality holds in (1.13), then Y is isomorphic to N . We do not have any general criterion on Y for existence of a mapping law such that Y is isomorphic to N . We will give an example for non-existence in Section 5.

We are interested in when we can take mapping laws such that the $h(N)$ approximates the $h(Y)$. Following Rosenblatt [14], we introduce the following:

Definition 1.5. A stationary Markov chain Y is called *p-uniform* if there exist a probability law ν on V and a family $\{\tau_x : x \in V\}$ of permutations of V such that

$$q_{x,y} = \nu(\tau_x(y)), \quad x, y \in V. \quad (1.14)$$

The second one of our main theorems is the following.

Theorem 1.6. *Let Y be a mixing Markov chain. Then the following assertions are equivalent:*

- (i) *There exists a sequence $\{\mu^{(n)} : n = 1, 2, \dots\}$ of mapping laws for Q with synchronizing support such that the $N^{(n)}$ corresponding to $\mu^{(n)}$ satisfies*

$$h(N^{(n)}) \rightarrow h(Y) \quad \text{as } n \rightarrow \infty. \quad (1.15)$$

- (ii) *Y is p-uniform.*

In particular, if $h(N) = h(Y)$ holds for N corresponding to some mapping law μ for Q with synchronizing support, then Y is necessarily p-uniform.

Theorem 1.6 will be proved in Section 4.

This paper is organized as follows. In Section 2, we introduce several notations to state the road coloring problem. Sections 3 and 4 are devoted to the proofs of Theorems 1.4 and 1.6, respectively. In Section 5, we give an example for Theorem 1.6.

2 Road colorings of a directed graph

Let $A = [A(y, x)]_{y, x \in V}$ be a $V \times V$ -matrix whose entries are non-negative integers. The pair (V, A) may be called a *directed graph*, where, for $x, y \in V$, the value $A(y, x)$ is regarded as the number of directed edges from x to y . The set V is called *the set of vertices* and the matrix A is called *the adjacency matrix*.

The graph (V, A) is called *of constant outdegree* if there exists a constant d such that

$$\sum_{y \in V} A(y, x) = d \quad \text{for all } x \in V. \quad (2.1)$$

In this case (V, A) is called *d-out*. The graph (V, A) is called *strongly connected* if, for any $x, y \in V$, there exists a positive integer $n = n(x, y)$ such that $A^n(y, x) \geq 1$. The graph (V, A) is called *aperiodic* if the greatest common divisor among $\{n \geq 1 : A^n(x, x) \geq 1\}$ is one for all $x \in V$. Note that (V, A) is both strongly connected and aperiodic if and only if there exists a positive integer r such that $A^r(y, x) \geq 1$ for all $x, y \in V$. We say that the graph (V, A) or the adjacency matrix A *satisfies the assumption (A)* if (V, A) is of constant outdegree, strongly connected and aperiodic.

Recall that Σ is the set of all mappings from V to itself. For $\sigma_1, \sigma_2 \in \Sigma$ and $x \in V$, we write $\sigma_2\sigma_1x$ simply for $\sigma_2(\sigma_1(x))$. The set Σ acts V in the following sense:

$$(\sigma_1\sigma_2)x = \sigma_1(\sigma_2x), \quad \sigma_1, \sigma_2 \in \Sigma, \quad x \in V. \quad (2.2)$$

The set $V = \{1, \dots, m\}$ may be identified with the set of standard basis $\{e_1, \dots, e_m\}$ of \mathbb{R}^m . An element $\sigma \in \Sigma$ may be identified with the 1-out adjacency matrix $\sigma = [\sigma(y, x)]_{y, x \in V}$ given as

$$\sigma = [\sigma e_1 \quad \dots \quad \sigma e_m]. \quad (2.3)$$

Under these identifications, we see that, for all $x, y \in V$,

$$\sigma(y, x) = 1 \quad \text{if and only if} \quad y = \sigma x. \quad (2.4)$$

Let (V, A) be a d -out directed graph. A family $\{\sigma^{(1)}, \dots, \sigma^{(d)}\}$ of elements of Σ (possibly with repeated elements) is called a *road coloring* of (V, A) if

$$A = \sigma^{(1)} + \dots + \sigma^{(d)}. \quad (2.5)$$

Each $\sigma^{(i)}$ is called a *road color*. Note that there exists at least one road coloring of (V, A) . Conversely, if we are given a family $\{\sigma^{(1)}, \dots, \sigma^{(d)}\}$ of elements of Σ (possibly with repeated elements), then it induces a unique d -out directed graph (V, A) given as (2.5).

Let Σ_0 be a subset of Σ . A sequence $s = (\sigma_p, \dots, \sigma_2, \sigma_1)$ of elements of Σ_0 is called a Σ_0 -word. For a Σ_0 -word $s = (\sigma_p, \dots, \sigma_2, \sigma_1)$, we write $\langle s \rangle$ for the product $\sigma_p \cdots \sigma_2 \sigma_1$. The following definition is a slight modification of Definition 1.3.

Definition 2.1. A road coloring $\Sigma_0 = \{\sigma^{(1)}, \dots, \sigma^{(d)}\}$ is called *synchronizing* if Σ_0 as a subset of Σ is synchronizing.

By this definition, we see that a road coloring $\Sigma_0 = \{\sigma^{(1)}, \dots, \sigma^{(d)}\}$ is synchronizing if and only if $\langle s \rangle V$ is a singleton for some Σ_0 -word s . If we express

$$s = (\sigma^{(i(p))}, \dots, \sigma^{(i(2))}, \sigma^{(i(1))}) \quad (2.6)$$

with some numbers $i(1), \dots, i(p) \in \{1, \dots, d\}$, the assertion “ $\langle s \rangle V$ is a singleton” may be stated in other words as follows: Those who walk in the graph (V, A) according to the colors $\sigma^{(i(1))}, \dots, \sigma^{(i(p))}$ in this order will lead to a common vertex, no matter where they started from.

Now we state the *road coloring theorem*.

Theorem 2.2 (Trahtman ([16])). *Suppose that the directed graph (V, A) satisfies the assumption (A). Then there exists a synchronizing road coloring of (V, A) .*

This was first conjectured in the case of no multiple directed edges by Adler–Goodwyn–Weiss [1] (see also [2]) in the context of the isomorphism problem of symbolic dynamics with common topological entropy. For related studies before Trahtman [16], see [7], [6] and [5].

3 Construction of a mapping law on a synchronizing road coloring

We need the following lemma.

Lemma 3.1. *Let Y be a stationary Markov chain with one-step transition probability matrix Q . Then there exists a mapping law μ for Q .*

Proof. First, we suppose that $q_{x,y}$ is a rational number for all $x, y \in V$. Then we may take an integer d sufficiently large so that $A(y, x) := q_{x,y}d$ is an integer for all $x, y \in V$. Then $A := [A(y, x)]_{x,y \in V}$ is the adjacency matrix of a d -out directed graph (V, A) ; in fact,

$$\sum_{y \in V} A(y, x) = d \sum_{y \in V} q_{x,y} = d. \quad (3.1)$$

Let $\{\sigma^{(1)}, \dots, \sigma^{(d)}\}$ be a road coloring of (V, A) and define

$$\mu(\sigma) = \frac{1}{d} \sharp(\{i = 1, \dots, d : \sigma^{(i)} = \sigma\}) \quad (3.2)$$

where $\sharp(\cdot)$ denotes the number of elements of the set indicated. Thus, for any $x, y \in V$, we see that

$$\sum_{\sigma \in \Sigma : y = \sigma x} \mu(\sigma) = \frac{1}{d} \sharp(\{i = 1, \dots, d : \sigma^{(i)}(y, x) = 1\}) = \frac{1}{d} A(y, x) = q_{x,y}, \quad (3.3)$$

which shows that μ is a mapping law for Q .

Second, we consider the general case. Let us take a sequence $\{Q^{(n)} : n = 1, 2, \dots\}$ of one-step transition probability matrices such that $q_{x,y}^{(n)}$ is a rational number for all n and $x, y \in V$ and that $q_{x,y}^{(n)} \rightarrow q_{x,y}$ as $n \rightarrow \infty$ for all $x, y \in V$. Then for any n there exists a mapping law $\mu^{(n)}$ for $Q^{(n)}$. Since Σ is a finite set, we can choose some subsequence $\{\mu^{(n(k))} : k = 1, 2, \dots\}$ and some probability law μ on Σ such that $\mu^{(n(k))}(\sigma) \rightarrow \mu(\sigma)$ as $k \rightarrow \infty$. This shows that μ is a mapping law for Q . The proof is now complete. \square

Now we proceed to prove Theorem 1.4.

Proof of Theorem 1.4. Let $Q = (q_{x,y})_{x,y \in V}$ be the one-step transition probability matrix for a mixing Markov chain Y .

First, we take an adjacency matrix A which is of constant outdegree such that

$$A(y, x) \begin{cases} \geq 1 & \text{if } q_{x,y} > 0, \\ = 0 & \text{if } q_{x,y} = 0. \end{cases} \quad (3.4)$$

For this, we introduce a subset $V \times V$ defined by

$$E = \{(x, y) \in V \times V : q_{x,y} > 0\}. \quad (3.5)$$

For each $x \in V$, we define the outdegree of E at x by

$$d(x) = \# \{(x, y) \in E : y \in V\} \quad (3.6)$$

and write $d = \max_{x \in V} d(x)$ for the maximum outdegree of E . For each $x \in V$, we may choose a site $\sigma(x) \in V$ so that $(x, \sigma(x)) \in E$. Now we may set

$$A(y, x) = \begin{cases} d - d(x) + 1 & \text{if } y = \sigma(x), \\ 1 & \text{if } y \neq \sigma(x) \text{ and } (x, y) \in E, \\ 0 & \text{otherwise.} \end{cases} \quad (3.7)$$

Then this $(A(y, x))_{x,y \in V}$ is as desired.

Since Y is a mixing Markov chain, there exists a positive integer r such that $q_{x,y}^r > 0$ for all $x, y \in V$. Hence we have $A^r(y, x) \geq 1$ for all $x, y \in V$; in fact, there exists a path $x = x_0, x_1, \dots, x_n = y$ such that $q_{x_{k-1}, x_k} > 0$ for $k = 1, 2, \dots, n$, which implies that $A(x_k, x_{k-1}) \geq 1$ for $k = 1, 2, \dots, n$. Thus we see that (V, A) satisfies the assumption **(A)**. Thus we may apply Theorem 2.2 to obtain a synchronizing road coloring $\{\sigma^{(1)}, \dots, \sigma^{(d)}\}$ of (V, A) . Define

$$\widehat{\mu}(\sigma) = \frac{1}{d} \# \{i = 1, \dots, d : \sigma^{(i)} = \sigma\}, \quad \sigma \in \Sigma \quad (3.8)$$

and define

$$\widehat{q}_{x,y} = \sum_{\sigma \in \Sigma : y = \sigma x} \widehat{\mu}(\sigma), \quad x, y \in V. \quad (3.9)$$

Then $\hat{\mu}$ is a mapping law for \hat{Q} and has synchronizing support. We also note that

$$\hat{q}_{x,y} = 0 \quad \text{if } (x,y) \notin E. \quad (3.10)$$

Let

$$\varepsilon = \min\{q_{x,y} : (x,y) \in E\} > 0. \quad (3.11)$$

If $\varepsilon = 1$, then we have $Q = \hat{Q}$, so that $\hat{\mu}$ is as desired. Let us assume that $\varepsilon < 1$. Define

$$Q^{(\varepsilon)} = \frac{1}{1-\varepsilon} (Q - \varepsilon \hat{Q}). \quad (3.12)$$

Then $Q^{(\varepsilon)} = (q_{x,y}^{(\varepsilon)})_{x,y \in V}$ is a one-step transition probability matrix of a stationary Markov chain. In fact, we see that

$$(1-\varepsilon)q_{x,y}^{(\varepsilon)} = q_{x,y} - \varepsilon \hat{q}_{x,y} \geq q_{x,y} - \varepsilon 1_{\{(x,y) \in E\}} \geq 0, \quad x, y \in V \quad (3.13)$$

and that

$$\sum_{y \in V} q_{x,y}^{(\varepsilon)} = \frac{1}{1-\varepsilon} \left(\sum_{y \in V} q_{x,y} - \varepsilon \sum_{y \in V} \hat{q}_{x,y} \right) = 1. \quad (3.14)$$

Now we apply Lemma 3.1 to obtain a mapping law $\mu^{(\varepsilon)}$ for $Q^{(\varepsilon)}$. Define

$$\mu = (1-\varepsilon)\mu^{(\varepsilon)} + \varepsilon \hat{\mu}. \quad (3.15)$$

Since $\mu^{(\varepsilon)}$ has synchronizing support, so does μ . For $x, y \in V$, we have

$$\sum_{\sigma \in \Sigma: y=\sigma x} \mu(\sigma) = (1-\varepsilon) \sum_{\sigma \in \Sigma: y=\sigma x} \mu^{(\varepsilon)}(\sigma) + \varepsilon \sum_{\sigma \in \Sigma: y=\sigma x} \hat{\mu}(\sigma). \quad (3.16)$$

$$= (1-\varepsilon)q_{x,y}^{(\varepsilon)} + \varepsilon \hat{q}_{x,y} = q_{x,y}, \quad (3.17)$$

which shows that μ is a mapping law for Q . The proof is now complete. \square

4 Approximate preservation of entropies

Let us prove Theorem 1.6.

Proof of Theorem 1.6. Let us prove that (i) implies (ii). Note that

$$h(Y) = - \sum_{x,y \in V} \lambda(x) q_{x,y} \log q_{x,y}, \quad (4.1)$$

$$h(N^{(n)}) = - \sum_{\sigma \in \Sigma} \mu^{(n)}(\sigma) \log \mu^{(n)}(\sigma). \quad (4.2)$$

Taking a subsequence if necessary, we may assume that there exists a probability law μ on Σ such that $\mu^{(n)}(\sigma) \rightarrow \mu(\sigma)$ for all $\sigma \in \Sigma$. Note that μ is a mapping law for Q but does not necessarily have synchronizing support. By the assumption (1.15), we see that

$$h(Y) = \lim_{n \rightarrow \infty} h(N^{(n)}) = - \sum_{\sigma \in \Sigma} \mu(\sigma) \log \mu(\sigma). \quad (4.3)$$

For $x, y \in V$, we set

$$\Sigma(y, x) = \{\sigma \in \Sigma : y = \sigma x\}, \quad (4.4)$$

so that we have

$$q_{x,y} = \sum_{\sigma \in \Sigma(y,x)} \mu(\sigma). \quad (4.5)$$

Hence we have

$$\mu(\sigma) \leq q_{x,y} \quad \text{whenever } \sigma \in \Sigma(y, x). \quad (4.6)$$

Since $t \mapsto \log t$ is increasing, we have

$$- \sum_{\sigma \in \Sigma(y,x)} \mu(\sigma) \log \mu(\sigma) \geq - \sum_{\sigma \in \Sigma(y,x)} \mu(\sigma) \log q_{x,y} = -q_{x,y} \log q_{x,y}. \quad (4.7)$$

Since $\bigcup_{y \in V} \Sigma(y, x) = \Sigma$, we have

$$h(Y) = - \sum_{y \in V} \sum_{\sigma \in \Sigma(y,x)} \mu(\sigma) \log \mu(\sigma) \geq q(x) \quad \text{for all } x \in V, \quad (4.8)$$

where we set

$$q(x) = - \sum_{y \in V} q_{x,y} \log q_{x,y}, \quad x \in V. \quad (4.9)$$

We take $\hat{x} \in V$ such that

$$q(\hat{x}) = \max_{x \in V} q(x). \quad (4.10)$$

Using (4.8) and (4.1), we have

$$q(\hat{x}) \leq h(Y) = \sum_{x \in V} \lambda(x) q(x) \leq q(\hat{x}). \quad (4.11)$$

Thus we see that the equalities hold in (4.11) and that $q(x) = q(\hat{x})$ for all $x \in V$. For any $x \in V$, we combine $h(N) = q(x)$ together with (4.7) and then obtain

$$- \sum_{\sigma \in \Sigma(y,x)} \mu(\sigma) \log \mu(\sigma) = -q_{x,y} \log q_{x,y}, \quad x, y \in V. \quad (4.12)$$

Combining this with (4.6), we obtain

$$\mu(\sigma) = q_{x,y} \quad \text{whenever } \sigma \in \Sigma(y, x). \quad (4.13)$$

Let $x_0 \in V$ be fixed and let $x \in V$. Since $\{\Sigma(y, x) : y \in V\}$ is a partition of Σ , we may choose a permutation τ_x of V so that

$$\Sigma(\tau_x(y), x) \cap \Sigma(y, x_0) \neq \emptyset, \quad y \in V. \quad (4.14)$$

This shows that

$$q_{x, \tau_x(y)} = q_{x_0, y}, \quad x, y \in V, \quad (4.15)$$

which implies p -uniformity of Y . The proof is now complete.

Let us prove that (ii) implies (i). Let $\{x_1, \dots, x_d\}$ be an enumeration of the support of the law ν in (1.14). For $i = 1, \dots, d$, we define

$$\sigma^{(i)}(y, x) = 1_{\{\tau_x(y) = x_i\}}. \quad (4.16)$$

For each $x \in V$, there exists a unique $y \in V$ such that $\sigma^{(i)}(y, x) = 1$, so that we have $\sigma^{(i)} \in \Sigma$. By (1.14), we obtain

$$q_{x,y} = \sum_{i=1}^d \sigma^{(i)}(y, x) \nu(x_i), \quad x, y \in V. \quad (4.17)$$

Let A be as in (3.4) and let Σ_1 be a synchronizing subset corresponding to some synchronizing road coloring of (V, A) . For sufficiently large integer n , we define a probability law $\mu^{(n)}$ on Σ by

$$\mu^{(n)}(\sigma) = \sum_{i: \sigma^{(i)} = \sigma} \left\{ \nu(x_i) - \frac{1}{nd} \right\} + \frac{1}{n|\Sigma_1|} 1_{\{\sigma \in \Sigma_1\}}. \quad (4.18)$$

Then it is obvious that $\mu^{(n)}$ is a mapping law for Q and has synchronizing support.

Let us verify the condition (1.15). On one hand, we have

$$h(N^{(n)}) \xrightarrow{n \rightarrow \infty} - \sum_{i=1}^d \nu(x_i) \log \nu(x_i). \quad (4.19)$$

On the other hand, we have

$$h(Y) = - \sum_{x,y \in V} \lambda(x) q_{x,y} \log q_{x,y} \quad (4.20)$$

$$= - \sum_{x,y \in V} \lambda(x) \sum_{i=1}^d \sigma^{(i)}(y, x) \nu(x_i) \log \nu(x_i) \quad (4.21)$$

$$= - \sum_{i=1}^d \left\{ \sum_{x,y \in V} \lambda(x) \sigma^{(i)}(y, x) \right\} \nu(x_i) \log \nu(x_i) \quad (4.22)$$

$$= - \sum_{i=1}^d \nu(x_i) \log \nu(x_i). \quad (4.23)$$

This shows (1.15). The proof is now complete. \square

5 An example

Let $V = \{1, 2\}$. Then $\Sigma = \{(12), (21), (11), (22)\}$ where

$$(ij) = \begin{bmatrix} 1 \mapsto i \\ 2 \mapsto j \end{bmatrix}, \quad i, j = 1, 2. \quad (5.1)$$

Let $0 < p < 1$ and consider a Markov chain Y with one-step transition probability given by

$$\begin{bmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \end{bmatrix} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}. \quad (5.2)$$

Then it is obvious that Y is a mixing Markov chain. Since

$$\begin{bmatrix} q_{1,1} \\ q_{2,1} \end{bmatrix} = \begin{bmatrix} q_{2,2} \\ q_{1,2} \end{bmatrix} = \begin{bmatrix} p \\ 1-p \end{bmatrix}, \quad (5.3)$$

we see that Y is p -uniform.

It is obvious that the stationary law is given as

$$\lambda(1) = \lambda(2) = \frac{1}{2}. \quad (5.4)$$

We now see that

$$h(Y) = \varphi(p) + \varphi(1-p) \quad (5.5)$$

where $\varphi(t) = -t \log t$.

If μ is a mapping law for Q , then we have

$$\mu(12) + \mu(11) = p, \quad \mu(21) + \mu(11) = 1-p. \quad (5.6)$$

From this, we see that there exists some ε with $0 \leq \varepsilon \leq \min\{p, 1-p\}$ such that

$$\varepsilon = \mu(11) = \mu(22), \quad \mu(12) = p - \varepsilon, \quad \mu(21) = 1 - p - \varepsilon. \quad (5.7)$$

Conversely, for any ε with $0 \leq \varepsilon \leq \min\{p, 1-p\}$, we may define $\mu = \mu^{(\varepsilon)}$ by equation (5.7) so that $\mu^{(\varepsilon)}$ is a mapping law for Q .

If $\mu^{(\varepsilon)}$ has synchronizing support, ε should be positive. Let $\{X^{(\varepsilon)}, N^{(\varepsilon)}\}$ be the $\mu^{(\varepsilon)}$ -random walk. We then see that

$$h(N^{(\varepsilon)}) = 2\varphi(\varepsilon) + \varphi(p - \varepsilon) + \varphi(1 - p - \varepsilon). \quad (5.8)$$

If $p = 1/2$, we see that $h(Y) = h(N^{(1/2)})$.

Suppose that $p \neq 1/2$. Then, by an easy computation, we may see that

$$h(Y) < h(N^{(\varepsilon)}) \quad (5.9)$$

for all ε with $0 < \varepsilon \leq \min\{p, 1-p\}$. However, it holds that $h(N^{(\varepsilon)}) \rightarrow h(Y)$ as $\varepsilon \rightarrow 0+$.

References

- [1] R. L. Adler, L. W. Goodwyn, and B. Weiss. Equivalence of topological Markov shifts. *Israel J. Math.*, 27(1):48–63, 1977.
- [2] R. L. Adler and B. Weiss. *Similarity of automorphisms of the torus*. Memoirs of the American Mathematical Society, No. 98. American Mathematical Society, Providence, R.I., 1970.
- [3] J. Akahori, C. Uenishi, and K. Yano. Stochastic equations on compact groups in discrete negative time. *Probab. Theory Related Fields*, 140(3-4):569–593, 2008.
- [4] P. Billingsley. *Ergodic theory and information*. Robert E. Krieger Publishing Co., Huntington, N.Y., 1978. Reprint of the 1965 original.
- [5] G. Budzban. Semigroups and the generalized road coloring problem. *Semigroup Forum*, 69(2):201–208, 2004.
- [6] G. Budzban and P. Feinsilver. Completely simple semigroups, Lie algebras, and the road coloring problem. *Semigroup Forum*, 74(2):206–226, 2007.
- [7] G. Budzban and A. Mukherjea. A semigroup approach to the road coloring problem. In *Probability on algebraic structures (Gainesville, FL, 1999)*, volume 261 of *Contemp. Math.*, pages 195–207. Amer. Math. Soc., Providence, RI, 2000.
- [8] N. A. Friedman and D. S. Ornstein. On isomorphism of weak Bernoulli transformations. *Advances in Math.*, 5:365–394 (1970), 1970.
- [9] O. Häggström. *Finite Markov chains and algorithmic applications*, volume 52 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2002.
- [10] T. Hirayama and K. Yano. Extremal solutions for stochastic equations indexed by negative integers and taking values in compact groups. *Stochastic Process. Appl.*, 120(8):1404–1423, 2010.
- [11] T. Hirayama and K. Yano. Strong solutions of Tsirelson’s equation in discrete time taking values in compact spaces with semigroup action. *Preprint, arXiv:1005.0038, submitted*, 2010.
- [12] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, 4:337–352 (1970), 1970.
- [13] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.

- [14] M. Rosenblatt. Stationary processes as shifts of functions of independent random variables. *J. Math. Mech.*, 8:665–681, 1959.
- [15] M. Rosenblatt. Stationary Markov chains and independent random variables. *J. Math. Mech.*, 9:945–949, 1960. Addendum: *ibid.* vol. 11, page 317, 1962.
- [16] A. N. Trahtman. The road coloring problem. *Israel J. Math.*, 172:51–60, 2009.
- [17] K. Yano. Random walk in a finite directed graph subject to a road coloring. *Preprint, arXiv:1005.0079*, 2010.
- [18] K. Yano and Y. Takahashi. Time evolution with and without remote past. *Sūrikaiseikikenkyūsho Kōkyūroku*, 1552:164–171, 2007. Recent Developments in Dynamical Systems (Kyoto, 2006).
- [19] K. Yano and M. Yor. Around Tsirelson’s equation, or: The evolution process may not explain everything. *Preprint, arXiv:0906.3442, submitted*, 2009.
- [20] M. Yor. Tsirel’son’s equation in discrete time. *Probab. Theory Related Fields*, 91(2):135–152, 1992.